



American Society for Gravitational and Space Research (ASGSR) 2016

GeneLab Data Systems (GLDS) Workshop

Peter B. Tran
Data Systems Lead
October 28th, 2016





- GLDS Phase 1 Key Capabilities (Operational)
 - Background History
 - Current Phase 1 C-Gene Key System Capabilities
- GLDS Phase 2 Key Capabilities (Work-In-Progress)
 - Introduction
 - Phase 2 X-Gene Key System Capabilities
- GLDS Tutorial (Phase 1 C-Gene)
- Backup - GLDS Metrics
 - Usage Metrics (Downloads, Page Views, etc.)



GLDS Phase 1 Key Capabilities (Operational)



Background History

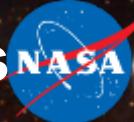
- GLDS Phase 1 (denoted as “C-Gene”) completed in April 2015 timeframe to implement and deploy as an interim solution as a data repository with simple search capability
- Initially imported 22 curated and V&V legacy data sets to Phase 1 C-Gene production system
- Phase 1 C-Gene is an operational system with routine maintenance for software enhancements (quarterly updates) & continuous data updates (6-8 weeks cycles)
- Publicly available at: <https://genelab-data.ndc.nasa.gov/>



- Since then, imported an additional 58 new & legacy curated, V&V data sets, including RR-1, ISS Microbial Observatory, zebrafish, etc., to Phase 1 C-Gene production system:
 - As of Release 1.0.14 deployed on 9/22/16, there is a total of **80 data sets** available for the general public
- Improved data importing script and optimize process to reduce site maintenance time (e.g., do not need to bring down production site during data updates)



- C-Gene Phase 1 C-Gene Operational System:
 - Software Enhancements:
 - Enhanced pagination (e.g., drop-down list of number of data sets displayed per page: defaults to 25 per page)
 - Added links to useful bioinformatics data mining & visualization tools (e.g., cBioPortal, Cistrome, GenePattern, GenomeSpace, ISA Tools, etc.)
 - Implemented a link to JSC Life Sciences Data Archive (LSDA) data request access portal based on certain metadata fields and values
 - Added GLDS accession numbers on the listing/browsing of data sets
 - Added a link to download the most recent GeneLab customized ISACreator configuration zip package for general public usage
 - Provided updates to genelab.nasa.gov contents (e.g., recent announcements, events, etc.)
 - Provided fixes to improve Look&Feel User Interfaces/Experience (UI/UX) (e.g., Mobile vs. Desktop screen sizes, intuitiveness)



- Leverage from existing NASA C3 core software platform and infrastructure (e.g., Don't reinvent the wheel!)
 - C3 successfully used by other NASA Programs/Projects (e.g., DASHlink, NASA Earth Exchange (NEX), and Applied Sciences Program Water Resources)
- Primary used as a data repository and displaying relevant metadata fields
- Browse and navigate user interfaces
- Download curated and V&V data sets as compressed zipped files
- Simple keyword search using regular expressions (e.g., pattern matching on certain fields like Title, unique GLDS Accession Numbers Study Description, Organisms, Contact Names, etc.)
- Links to relevant publications (if available)



<http://genelab.nasa.gov>



GeneLab

Open Science
for Exploration

Home →
About
Announcements
Events
Data Repository
Contact Us
Join Us!

→ EXPLORE OUR
DATA REPOSITORY

▶ Watch our story —

This is GeneLab



Cutting-edge biology experiments
conducted in space. *Discoveries made.*
Results analyzed on big data
platforms. Research shared. *This is*
science amplified. What will you
discover?

Sign up today —

Become a member of
the GeneLab network
and discover science
that goes beyond.

JOIN US





NASA GeneLab Data Repository

NASA GeneLab will expand scientists' access to experiments that explore the molecular response of terrestrial biology to spaceflight environments. The vast amounts of raw data generated by experiments aboard the International Space Station will be made available to a worldwide community of scientists and computational researchers.

This data repository holds NASA's space biology datasets.

For more information about NASA GeneLab, please refer to our overview website at genelab.nasa.gov. Another source of information is the [GeneLab Strategic Plan](#).

Study Metadata

Study data hosted in the repository include a metadata definition file that is formatted according the ISA-Tab specification, which is viewable using the ISA Tools ISACreator program using a customized GeneLab configuration. Consumers of the metadata files should load this configuration, [ISA configuration files](#), into the ISACreator program prior to loading the ISA-Tab file set, in order to view the study metadata completely and correctly.

GeneLab Sponsors

NASA GeneLab is sponsored by the [NASA Space Life and Physical Sciences Research and Applications Division \(SLPSRA\)](#) and the [NASA International Space Station Program](#).

Contributing Datasets to GeneLab

Please refer to our [data submission process](#) to contribute your data to NASA GeneLab.





Simple Navigating and Browsing

Microbial Observatory (ISS-MO): Indoor microbiome study of the International Space Station surfaces



GLDS-69

Microbial Observatory (ISS-MO): Draft Genome Sequence of two *Aspergillus fumigatus* Strains Isolated from the International Space Station



GLDS-68

Microbial Observatory (ISS-MO): Study of BSL-2 bacterial isolates from the International Space Station



GLDS-67

Microbial Observatory (ISS-MO): Antimicrobial resistance genes



GLDS-66

Simple Metadata Display

Microbial Observatory (ISS-MO): Indoor microbiome study of the International Space Station surfaces



2 datasets available for download here:

Study Metadata Files

Metagenomic Sequencing Assay Data Files

| | | | |
|--------------------------|--|------------------------|--------------|
| GeneLab Accession Number | GLDS-69 | | |
| Source Accession Number | N/A | | |
| Contacts | Name | Role | Organization |
| | Kasthuri Venkateswaran | Principal Investigator | NASA |
| | kjvenkat@jpl.nasa.gov | | |
| Submission Date | 03-May-2016 | | |
| Public Release Date | 07-Jul-2016 | | |
| Study Description | Presented here is the environmental microbiome study of the International Space Station surfaces. The environmental samples were collected with the polyester wipes from eight different locations in the ISS during two consecutive sampling sessions (three months apart). The specific objective was to unveil the pool of genes for each location during two separate sessions to learn of functional and metabolic diversity of microorganisms in the ISS. The International Space Station (ISS) as a closed built environment has its own environmental microbiome which is shaped by microgravity, radiation, and limited human presence. The microbial diversity associated with ISS environmental surfaces was investigated during this study. Polyester wipes and contact slides were used for sampling of eight various surface locations on the ISS at different time periods. The samples were retrieved and analyzed immediately upon the return to the Earth (via Soyuz TMA-14M or Dragon capsule from SpaceX). After surface sample collection, contact slides containing nutrient media for the growth of bacteria and fungi were incubated at 25°C. The polyester wipes were processed to measure microbial burden (R2A, Blood Agar and Potato Dextrose Agar) and recover cultivable bacteria as well as fungi. Subsequently, viable microbial burden was assessed using Adenosine Triphosphate (ATP) assay, and quantitative polymerase chain reaction (PCR) methods after propidium monoazide (PMA) treatment. The 16S-tag and metagenome analyses were used to elucidate viable microbial diversity. The cultivable bacterial population yield from the polyester wipes was very high (5 to 7-logs) when compared with the contact slides (102 to 103 CFU/m2). The PMA-qPCR analysis showed considerable variation of viable bacterial population (105 to 109 16S rDNA gene copies/m2) among locations sampled. Unlike contact slides, polyester wipes cover much larger sample surface (~1 m2) and produce much more reliable results of the microbial diversity of the ISS covering both cultivable and non-cultivable species. The cultivable, total, and viable microbial diversity was determined utilizing state-of-the-art molecular techniques. The implementation of the PMA assay before DNA extraction allowed distinguishing viable microorganisms, which is crucial for determining their role to the crew health, the ISS maintenance and the general knowledge of the closed environmentally controlled built systems. | | |
| Organisms | cellular organisms | | |
| Study Design Factor(s) | Factor | Ontology: Concept | |
| | spaceflight | Space Flight | |



GLDS Phase 2 Key Capabilities (Work-In-Progress)



- GLDS Phase 2 (denoted as “X-Gene”) **Work-In-Progress**:
 - Focus on integrated full-text search & data federation key capabilities (e.g., data integration/fusion and mash-up)
 - Target Phase 2 completion (Release 2.0) with fully implemented capabilities available worldwide in Sept. 2017
 - Recently completed Release 1.1 (Phase 2 Iteration 1) on 9/14/16 as initial proof-of-concept integrated search + data federation with NIH/NCBI’s GEO database
 - Starting on Release 1.2 (Phase 2 Iteration 2) with target completion in Jan. 2017 with integrated search + data federation enhancements (e.g., add new external bioinformatics data sources)
 - Starting to solicit internal user feedback (e.g., looking for possible internal participants)



- A Realigned Trade Study:
 - The basis for new X-Gene system (Phase 2 and beyond)
 - Will eventually replace existing Phase 1 C-Gene operational system built on top of ARC-TI Center for Cross-discipline Collaboration (C3) platform
 - Takes lessons learned from FY15 trade study and will not start from scratch:
 - Select top 7 software platforms from FY15 with appropriate rationales + 3 new platforms not evaluated previously, such as:
 - NSF/University of Arizona's iPlant/CyVerse, Broad Institute's (Harvard/MIT) GenomeSpace, and NASA's Athena platform
 - Total of 10 platforms to evaluate & assess for FY16



- Reached out to 78 external participants from:
 - Science Council
 - Academia
 - Other Government Agencies (e.g., Veterans Affairs Dept., Lawrence Berkeley/Livermore National Labs, DoD, etc.)
 - NASA Space Biology, HRP, and other PIs
 - GeneLab Stakeholders (e.g., SLPS & ISSP)
 - Other NASA Centers (e.g., JSC, KSC, HQ, etc.)
- Received a total of 16 participants' Excel input submissions:
 - 7 external participants (e.g., representatives from VA Medical Center, U.S. Army Center for Environmental Health Research, Space Biology/HRP PIs, Science Council, Duke University, etc.) compared to last FY15 had 2 GL stakeholders as “external” participants
 - 9 internal GeneLab participants compared to FY15 had only 7 participants



- Statistical analysis on the following:
 - RAW Average (Mean) score across all 16 participants (0 through 5 ratings for each criteria) for each software platform (total of 10 platforms evaluated); Hence a total of $16 \times 30 \times 10 = 4,800$ data points.
 - Weighted Average (Mean) score across all participants for each platform with reassigned weight factor (e.g., all 30 criteria weight % must total 100%)
 - RAW Median score (e.g., what's the middle value)
 - RAW Mode score (e.g., most frequent occurred value)
 - Sampling vs. Population Variances (e.g., deviation from the mean)
 - Standard Deviations (e.g., measures the amount of variation)
- Other important factors & considerations:
 - Software Maintainability & Complexity
 - Count total physical Source Lines of Code (SLOC) to measure codebase size
 - Different popular programming languages used in bioinformatics field (e.g., Java, JavaScript, Python, Perl, R, etc.)
 - Number of files & comments in codebase
 - Number of source code repositories (e.g., GitHub and/or other locations)



| | FY16 Trade Study (Ranked) | RAW Avg Score | Weighted Avg Score | RAW Median Score | RAW Mode Score | Sample Variance | Population Variance | Sample Std. Deviation | Population Std. Deviation |
|---|---------------------------|---------------|--------------------|------------------|----------------|-----------------|---------------------|-----------------------|---------------------------|
| 1 | GenomeSpace | 3.92 | 4.02 | 4.00 | 4.00 | 0.46 | 0.44 | 0.68 | 0.67 |
| 2 | KBase | 3.97 | 4.00 | 4.14 | 4.00 | 0.45 | 0.44 | 0.67 | 0.66 |
| 3 | CyVerse | 3.91 | 3.98 | 4.09 | 4.47 | 0.55 | 0.53 | 0.74 | 0.73 |
| 4 | Arvados | 3.67 | 3.78 | 3.76 | 3.62 | 0.41 | 0.40 | 0.64 | 0.63 |
| 5 | Globus | 3.58 | 3.64 | 3.79 | 4.00 | 0.46 | 0.44 | 0.68 | 0.64 |



- Top 5 highest ranked platform candidates by weighted average scores are:
 1. GenomeSpace from Broad Institute (Harvard/MIT)
 2. KBase from DOE's Office of Science
 3. CyVerse (formerly iPlant) from NSF & University of Arizona
 4. Arvados from Harvard University
 5. Globus from Computation Institute (University of Chicago & Argonne National Lab)



- Based on FY16 trade study results, the platform recommendation is...GenomeSpace!
 - Will meet programmatic maintainability of source codebase within resource guidelines
 - Predominately written in Java and JavaScript programming languages (current GLDS team has expertise!)
 - Provides modular system architecture & design for extensibility, scalability, flexibility, and performance using Service-Oriented Architecture (SOA), similar to GLDS generic reference system architecture
 - Best option for “T-Gene” and possibly “NRA Gene” capabilities
- Pivoting to new recommended software platform GenomSpace coming out of FY16 realigned trade study (e.g., migration plan for both software + data in Phase 2)



- GeneLab identified 3 key strategic external bioinformatics data sources for Phase 2 integration:
 - NIH/NCBI's GEO database (Gene Expression Omnibus)
 - EBI's PRIDE database (Proteomics)
 - MG-RAST database (Metagenomics)
- Recommendations from GeneLab Steering Committee in July 2016 to possibly include:
 - Nuclear Receptor Signaling Atlas (NURSA)
 - NIH's Human Microbiome Project Data Analysis & Coordination Center (HMPDACC)
 - NIH/NCBI's database of Genotypes & Phenotypes (dbGaP)



- GLDS Phase 2 X-Gene (Integrated Full-Text Search & Data Federation)
 - Release 1.1 deployed on 9/14/16
 - Prototype of integrated full-text search & data federation/integration between GeneLab & NIH/NBCI's GEO databases (e.g., Google-like search engine)
 - Establish core Common Metadata Model (CMM) for easier data federation/integration to other bioinformatics databases
 - Customizable search ranking/relevancy algorithm
 - Provides Medical Subject Headings (MeSH) as a thesaurus for “smart” searching on metadata fields
 - Frequent GeneLab specific keyword search terms lookup
 - Data visualization (GLDS Organisms vs. Assay Types)
 - Automatic scripts to pull GEO data changes on nightly basis



GeneLab

Open Science for Exploration

Phase 2 X-Gene Release 1.1 Search and Federation



GeneLab

Open Science
for Exploration

Google-like, full-text search with data federation/integration to NIH GEO database (over 80K GSE metadata records ingested)

Home Repository Data Data Mining Tools Submit Data Contact Us

✓ All
GeneLab
NIH GEO



Search results for: **mouse liver STS**

Total Search Results Found: **12440**

Sort by Relevance 25

1 2 3 4 Next >>

STS-135 Liver Transcriptomics

<http://genelab-app-1-stage-pubvpc.nasawestprime.com/genelab/acquisition/GLDS-25/>



Mice were flown onboard STS-135 and returned to Earth for analysis. Livers were collected within 3-4 hours of landing and snap frozen in liquid nitrogen. Samples were shipped to UCI Genomics High Throughput Facility for analysis.

Organism: *Mus musculus* Factor: Space Flight Assay Type: DNA microarray Accession: GLDS-25
PI/Contact: Michael J Pecaut, Xiao W... Release/Publication Date: Dec-31-1969

Space Transportation System (**STS**) Jul 21 2011 Jul 07 2011 **STS-135 liver** CBTM-3 (Biospecimen Sharing Program),transformation , data transformation NASA **STS-135 Liver** Transcriptomics Space Flight 0000000000 ht,ciences/cbtm-3-**sts-135** Biological and metabolic response in **STS-135** space-flown **mouse** skin. , Changes,Changes in **mouse** thymus and spleen after return from the **STS-135** mission in space. , Spaceflight environment,(ARC) Mice were flown onboard **STS-135** and returned to Earth for analysis. **Livers** were collected within 3-4

Murine liver tissues: WT and Wip1 KO mice after Partial Hepatectomy at 24h and 36 h.

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE59023>

Transcriptional profiling of mouse liver tissues comparing Wild type liver tissues with Wip1 KO mice liver tissues after Partial Hepatectomy at 24h and 36 h.

Organism: *Mus musculus* Accession: GSE59023 PI/Contact: Hu Wang Release/Publication Date: Jul-03-2014

GPL17261 Wang Hu Murine **liver** tissues: WT and Wip1 KO mice after Partial **Hepatectomy** at 24h and 36 h. 1404370800,profiling of **mouse liver** tissues comparing Wild type **liver** tissues with Wip1 KO mice **liver** tissues after,after Partial **Hepatectomy** at 24h and 36 h. *Mus musculus*

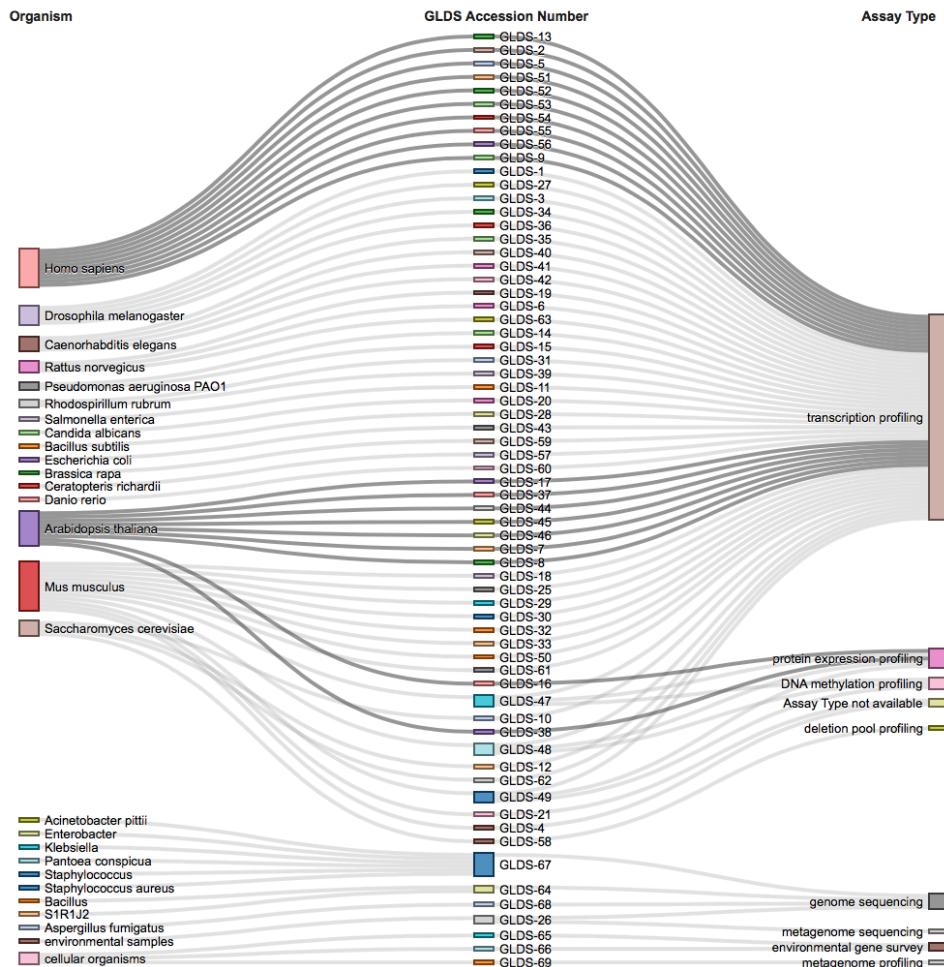


Home Repository Data Data Mining Tools Submit Data Contact Us

Search Data



GeneLab





GLDS Tutorial (Phase 1 C-Gene)

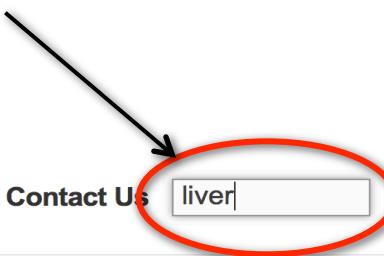


Home Repository Data Data Mining Tools Submit Data Contact Us

liver



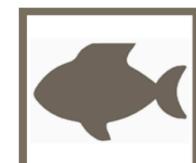
Enter keyword “liver” on search textbox



Page 1 of 4 (Total Studies: 80) [Next >](#)

Studies Per Page:

25



GLDS-83

Comparative Transcriptomic Analysis of Adult Medaka Tissues Sampled after Adaptation to a Space Environment

| Organisms | Factors | Assay Types | Release Date | Description |
|-----------------|-------------|-------------------------|--------------|---|
| Oryzias latipes | Spaceflight | transcription profiling | 15-Oct-2015 | To understand how humans adapt to space environments, many experiments can be conducted on astronauts while they work aboard the Space Shuttle or the International Space Station (ISS). We also need animal experiments that can apply to human... |



GLDS-81

Bacillus subtilis strains at low-pressure: 5 kPa versus 101 kPa growth

| Organisms | Factors | Assay Types | Release Date | Description |
|-------------------|-----------------|-------------------------|--------------|---|
| Bacillus subtilis | strain pressure | transcription profiling | 14-Mar-2014 | Comparing the transcriptional responses of Bacillus subtilis strains WN624 and WN1106 at 5 kPa and 101 kPa. WN1106 is a 5 kPa-evolved strain with increased fitness compared to ancestor-WN624 strain at 5 kPa. This experiment probed the diffe... |



5 search results for "liver"



GLDS-63

Transcription profiling of rat response to changes in developmental stage - 3 types of tissue, 3 gravity conditions, 2 developmental conditions

| Organisms | Factors | Assay Types | Release Date | Description |
|-------------------|--|----------------------------|--------------|--|
| Rattus norvegicus | tissue microgravity developmental condition hypergravity | transcription profiling | 27-Nov-2009 | Transcriptional crosstalk between mammary gland, liver and adipose tissue Experiment Overall Design: Pregnant and Lactating rats exposed to 3 gravity conditions |

Returned 5 search results
for keyword "liver"



GLDS-49

Multi-omic investigations of mouse liver subjected to simulated spaceflight freezing and storage protocols

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|--|---|--------------|---|
| Mus musculus | freezing profile tissue storage time | DNA methylation profiling protein expression profiling transcription profiling | Feb-04-2016 | This study compares standard laboratory protocols for tissue freezing and storage with a simulation of the delayed processing of liver specimens and long-term storage protocols used during the Rodent Research-1 (RR-1) payload. Liver samples... |



GLDS-48

Rodent Research-1 (RR1) NASA Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|--|---|--------------|---|
| Mus musculus | gravitation dissection condition | DNA methylation profiling transcription profiling protein expression profiling | Dec-15-2015 | RR-1 is a validation flight to evaluate the hardware, operational and science capabilities of the Rodent Research Project on the ISS. RNA, DNA, and protein were purified from 14 livers from RR-1 mice (female, C57Bl6/J, 16wk old at time of l... |



GLDS-47

Rodent Research-1 (RR1) National Lab Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|---------------------------------------|---|--------------|--|
| Mus musculus | Gravitation Dissection Timeline | DNA methylation profiling protein expression profiling transcription profiling | 15-Dec-2015 | The Rodent Research-1 National Lab (RR-1 CASIS) experiment was performed to study the effect of microgravity on muscle wasting. RNA, DNA, and protein were purified from nine RR-1 CASIS (the Center for the Advancement of Science in Space) l... |



GLDS-25

STS-135 Liver Transcriptomics

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|--------------|----------------------------|--------------|---|
| Mus musculus | Space Flight | transcription profiling | Oct 29 2015 | Mice were flown onboard STS-135 and returned to Earth for analysis. Livers were collected within 3-4 hours of landing and snap frozen in liquid nitrogen. |



GLDS-63

Transcription profiling of rat response to changes in developmental stage - 3 types of tissue, 3 gravity conditions, 2 developmental conditions

| Organisms | Factors | Assay Types | Release Date | Description |
|-------------------|---|-------------------------|--------------|---|
| Rattus norvegicus | tissue microgravity developmental condition hypergravity | transcription profiling | 27-Nov-2009 | Transcriptional crosstalk between mammary gland, liver and adipose tissue Experiment Overall Design: Pregnant and Lactating rats exposed to 3 gravity conditions |

Click on title link for more information



GLDS-49

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|---|--|--------------|---|
| Mus musculus | freezing profile tissue storage time | DNA methylation profiling protein expression profiling transcription profiling | Feb-04-2016 | This study compares standard laboratory protocols for tissue freezing and storage with a simulation of the delayed processing of liver specimens and long-term storage protocols used during the Rodent Research-1 (RR-1) payload. Liver samples... |



GLDS-48

Rodent Research-1 (RR1) NASA Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data

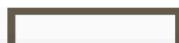
| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|-------------------------------------|--|--------------|--|
| Mus musculus | gravitation dissection condition | DNA methylation profiling transcription profiling protein expression profiling | Dec-15-2015 | RR-1 is a validation flight to evaluate the hardware, operational and science capabilities of the Rodent Research Project on the ISS. RNA, DNA, and protein were purified from 14 livers from RR-1 mice (female, C57Bl6J, 16wk old at time of l... |



GLDS-47

Rodent Research-1 (RR1) National Lab Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data

| Organisms | Factors | Assay Types | Release Date | Description |
|--------------|------------------------------------|--|--------------|--|
| Mus musculus | Gravitation Dissection Timeline | DNA methylation profiling protein expression profiling transcription profiling | 15-Dec-2015 | The Rodent Research-1 National Lab (RR-1 CASSIS) experiment was performed to study the effect of microgravity on muscle wasting. RNA, DNA, and protein were purified from nine RR-1 CASSIS (the Center for the Advancement of Science in Space) l... |



STS-135 Liver Transcriptomics



Rodent Research-1 (RR1) NASA Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data



4 datasets available for download here:

Proteomics Data Files
Study Metadata Files
Bisulfite Sequencing Data Files
RNA-Seq Data Files

Option to download
relevant data sets

| | | | |
|--------------------------|--|--|--|
| GeneLab Accession Number | GLDS-48 | | |
| Source Accession Number | | | |
| Contacts | Name | Role | Organization |
| | Ruth Globus | RR1 Project Scientist | NASA ARC |
| Submission Date | | | |
| Public Release Date | Dec-15-2015 | | |
| Study Description | RR-1 is a validation flight to evaluate the hardware, operational and science capabilities of the Rodent Research Project on the ISS. RNA, DNA, and protein were purified from 14 livers from RR-1 mice (female, C57BL/6J, 16wk old at time of launch), including seven from the Flight group and seven from the Ground Control group. From each group, two liver samples were collected and frozen immediately after euthanasia (Flight mice dissected on-orbit after total 37 days after launch, Samples FLT-2,3 and corresponding Ground Control samples GC-21,22). An additional five samples from each group were collected from frozen carcasses dissected post-flight (Samples FLT-26,27,28,29,30 and corresponding Ground Control samples GC-36,37,38,39,40). RNA-Seq, whole genome BS-Seq (bisulfite sequencing) and proteomic expression profiling were performed. | | |
| Organisms | Mus musculus | | |
| Study Design Factor(s) | Factor | Ontology: Concept | |
| | gravitation | Gravitation | |
| | dissection condition | dissection | |
| Assay(s) | Assay Type | Device Type | Device Platform |
| | DNA methylation profiling | nucleotide sequencing | Illumina |
| | transcription profiling | nucleotide sequencing | Illumina |
| | protein expression profiling | mass spectrometry | LTQ Orbitrap Velos (Thermo Scientific) |
| Project | Project Identifier | RR-1 | |
| | Project Link | https://lsda.jsc.nasa.gov/scripts/experiment/exper.aspx?exp_index=13380 | |
| | Project Type | Flight Study | |
| | Flight Program | International Space Station (ISS) | |
| | Experiment Platform | AEM-X Habitat | |
| | Space Program | NASA | |
| | Managing NASA Center | Ames Research Center (ARC) | |
| | Funding Source | This investigation was funded by the NASA Space Biology Program Office, Space Life and Physical Sciences Research and Applications Division, and additional funding from the | |



Datasets for Study: Rodent Research-1 (RR1) NASA Validation Flight: Mouse liver transcriptomic, proteomic, and epigenomic data

Total Data Volume: 367.2 GB



RNA-Seq Data Files

Raw sequencing files are in FASTQ format.

File Download Links:

[GLDS-40_transcriptomics_N-GC-37.tar.gz](#) (1.0 GB)
[GLDS-48_transcriptomics_N-GC-36.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-GC-22.tar.gz](#) (1.2 GB)
[GLDS-48_transcriptomics_N-GC-21.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-FLT-3_.tar.gz](#) (960.4 MB)
[GLDS-48_transcriptomics_N-FLT-30.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-FLT-2_.tar.gz](#) (1.3 GB)
[GLDS-48_transcriptomics_N-FLT-28.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-FLT-26.tar.gz](#) (1.2 GB)
[GLDS-48_transcriptomics_N-FLT-27.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-FLT-25.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_N-GC-38.tar.gz](#) (1.3 GB)
[GLDS-48_transcriptomics_N-GC-39.tar.gz](#) (1.0 GB)
[GLDS-48_transcriptomics_N-GC-40.tar.gz](#) (1.1 GB)
[GLDS-48_transcriptomics_RR1-NASA.md5sum](#) (2.8 KB)

**Options to download
either RNA-Seq or
Proteomics data files**



Proteomics Data Files

LC-MS3 Proteomics Data. Compressed collection of raw or processed data files and quality report associated with this study. Formats are platform specific. Please view the associated ISA-TAB metadata files to get formatting details.

File Download Links:

[GLDS-48_proteomics_RR1-NASA.md5sum](#) (1.5 KB)
[GLDS-48_proteomics_RR1-NASA.raw.tar.gz](#) (11.1 GB)
[GLDS-48_proteomics_RR1-NASA.processed.tar.gz](#) (8.4 GB)



Study Metadata Files

ISA-Tab format file. Tab delimited Investigation, Study, and Assay level metadata. See ISA-tools.org for format details and software.

File Download Links:

[GLDS-48_metadata_RR1-NASA-ISA.zip \(9.0 KB\)](#)

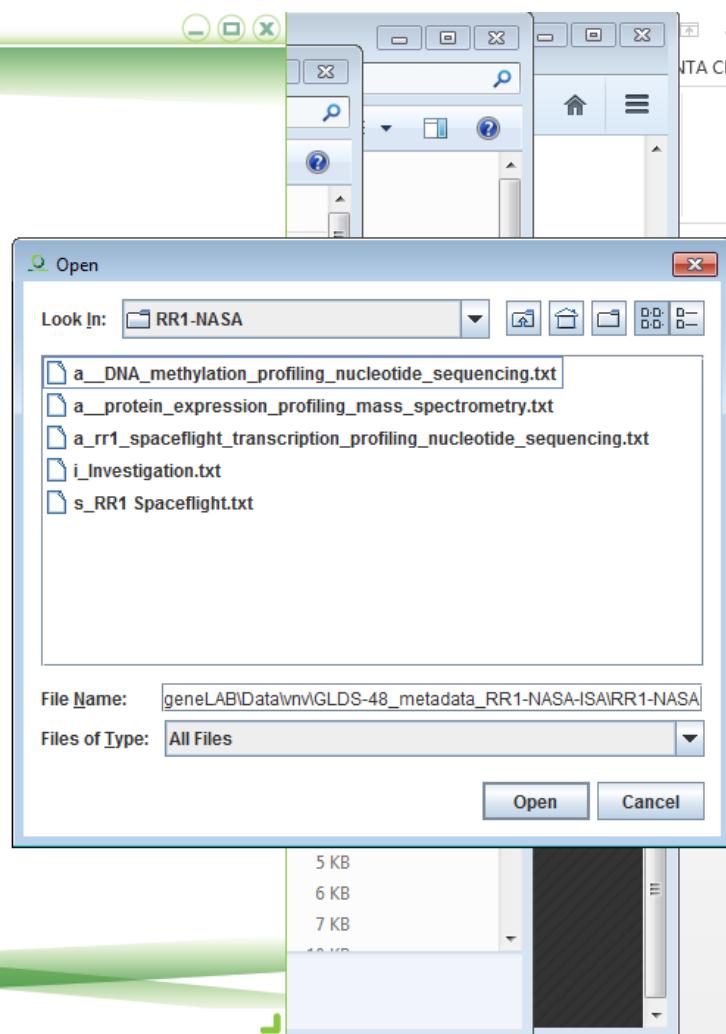
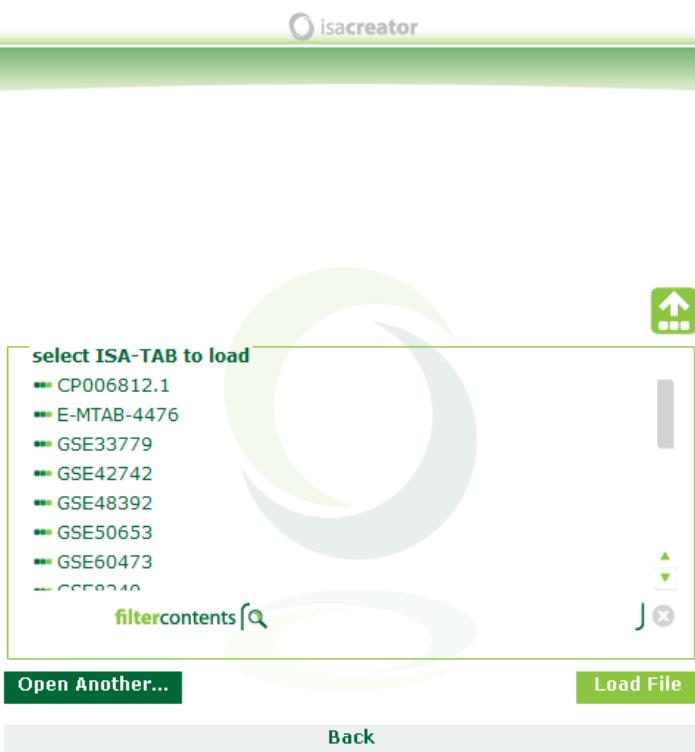
**Click link to download
ISA-Tab formatted
metadata zip package**

<https://genelab-data.ndc.nasa.gov/>

Study Metadata

Study data hosted in the repository include a metadata definition file that is formatted according the ISA-Tab specification, which is viewable using the ISA Tools ISACreator program using a customized GeneLab configuration. Consumers of the metadata files should load this configuration, [ISA configuration files](#), into the ISACreator program prior to loading the ISA-Tab file set, in order to view the study metadata completely and correctly.

Click the “ISA configuration files” link on GLDS repository splash page to download GeneLab customized ISACreator configuration zip package





isatab™ overview

file study view utilities options help

isacreator

Sample Definitions

| Row No. | Source Name | Characteristics[age] | Unit | Characteristics[feeding] | Characteristics[strain] | Characteristics[MESH:Sex] | Factor Value[gravitation] | Factor Value[temperature] |
|---------|-------------|----------------------|----------|--------------------------|-------------------------|---------------------------|---------------------------|---------------------------|
| 1 | GC 21 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | upon |
| 2 | GC 22 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | upon |
| 3 | FLT 2 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | upon |
| 4 | FLT 3 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | upon |
| 5 | GC 36 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | carca |
| 6 | GC 37 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | carca |
| 7 | GC 38 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | carca |
| 8 | GC 39 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | carca |
| 9 | GC 40 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | ground control | carca |
| 10 | FLT 25 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | carca |
| 11 | FLT 26 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | carca |
| 12 | FLT 27 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | carca |
| 13 | FLT 28 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | carca |
| 14 | FLT 30 | | 21 weeks | ad libitum | C57/BL6 | EFO:female | spaceflight | carca |

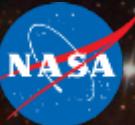
information

ontology term information

liver

source ref: EFO

View Resource



isatab overview

RR1-NASA

s_RR1_Spaceflight.txt

a__DNA_methylation_p

a_rr1_spaceflight_trans

a__protein_expression

Comment[Export]

Raw Data File

Protocol REF

Normalization Name

Data Transformation Name

Derive

Assay measuring transcription profiling using OBI:nucleotide sequencing

| | Raw Data File | Protocol REF | Normalization Name | Data Transformation Name | Derive |
|----|----------------------------------|--------------|--------------------|--------------------------|--------|
| no | *transcriptomics_N-GC-21.tar.gz | | | | |
| no | *transcriptomics_N-GC-22.tar.gz | | | | |
| no | *transcriptomics_N-FLT-2.tar.gz | | | | |
| no | *transcriptomics_N-FLT-3.tar.gz | | | | |
| no | *transcriptomics_N-GC-36.tar.gz | | | | |
| no | *transcriptomics_N-GC-37.tar.gz | | | | |
| no | *transcriptomics_N-GC-38.tar.gz | | | | |
| no | *transcriptomics_N-GC-39.tar.gz | | | | |
| no | *transcriptomics_N-GC-40.tar.gz | | | | |
| no | *transcriptomics_N-FLT-25.tar.gz | | | | |
| no | *transcriptomics_N-FLT-26.tar.gz | | | | |
| no | *transcriptomics_N-FLT-27.tar.gz | + | | | |
| no | *transcriptomics_N-FLT-28.tar.gz | | | | |
| no | *transcriptomics_N-FLT-30.tar.gz | | | | |

information

Assay



Citing NASA GeneLab Data Sets

Distribution of GeneLab data sets is funded by NASA's Space Life and Physical Sciences Research and Applications and International Space Station Research Office.

www.nasa.gov/directorates/heo/slprsa/

www.nasa.gov/mission_pages/station/main

These data are provided under NASA's open access policy. To acknowledge the scientists who have provided data, we request that you include a bibliographic citation to all GeneLab-provided data that you use in your publications. Such citations will help others find the products and see how they have been used. Citation information is provided below. If you have questions about how to cite GeneLab data or services, please contact GeneLab support at arc-dl-genelab-datasystems@mail.nasa.gov?subject=GeneLab%20Citation.

Acknowledgement:

A general statement crediting NASA GeneLab for data, assistance, and/or review. This statement is included in a paragraph at the end of an article, before the reference list.

Example Acknowledgment for GeneLab data:

"GeneLab data are courtesy of the NASA GeneLab Data Repository (<https://genelab-data.ndc.nasa.gov/genelab/projects/>)."

Reference Citation:

Author (or NASA GeneLab as producer), product title, publisher (NASA GeneLab), and release date of the data product which is being credited.

Example Reference Citations:

For a Dataset:

Doe JQ, Smith RS. 2010. Effects of spaceflight on *D. melanogaster* gene expression. Dataset. Available on-line [<https://genelab-data.ndc.nasa.gov/genelab/acquisition/GLDS-XXX>] from NASA GeneLab, NASA Ames Research Center, Moffett Field, CA.

For a Citation on a Website:

"*D. melanogaster* gene expression dataset (2015) obtained from <https://genelab-data.ndc.nasa.gov/genelab/projects/>, maintained by NASA GeneLab, NASA Ames Research Center, Moffett Field, CA 94035. "

Reprints

If possible, please e-mail arc-dl-genelab-datasystems@mail.nasa.gov?subject=GeneLab%20Citation or send via postal mail reprints/citations of papers or oral presentations based on data obtained from the GeneLab Data Repository. Your input will help us to stay informed as to how the data are used.



Thank You!

Contact Us

genelab-outreach@lists.nasa.gov



Backup Slides



GLDS Metrics

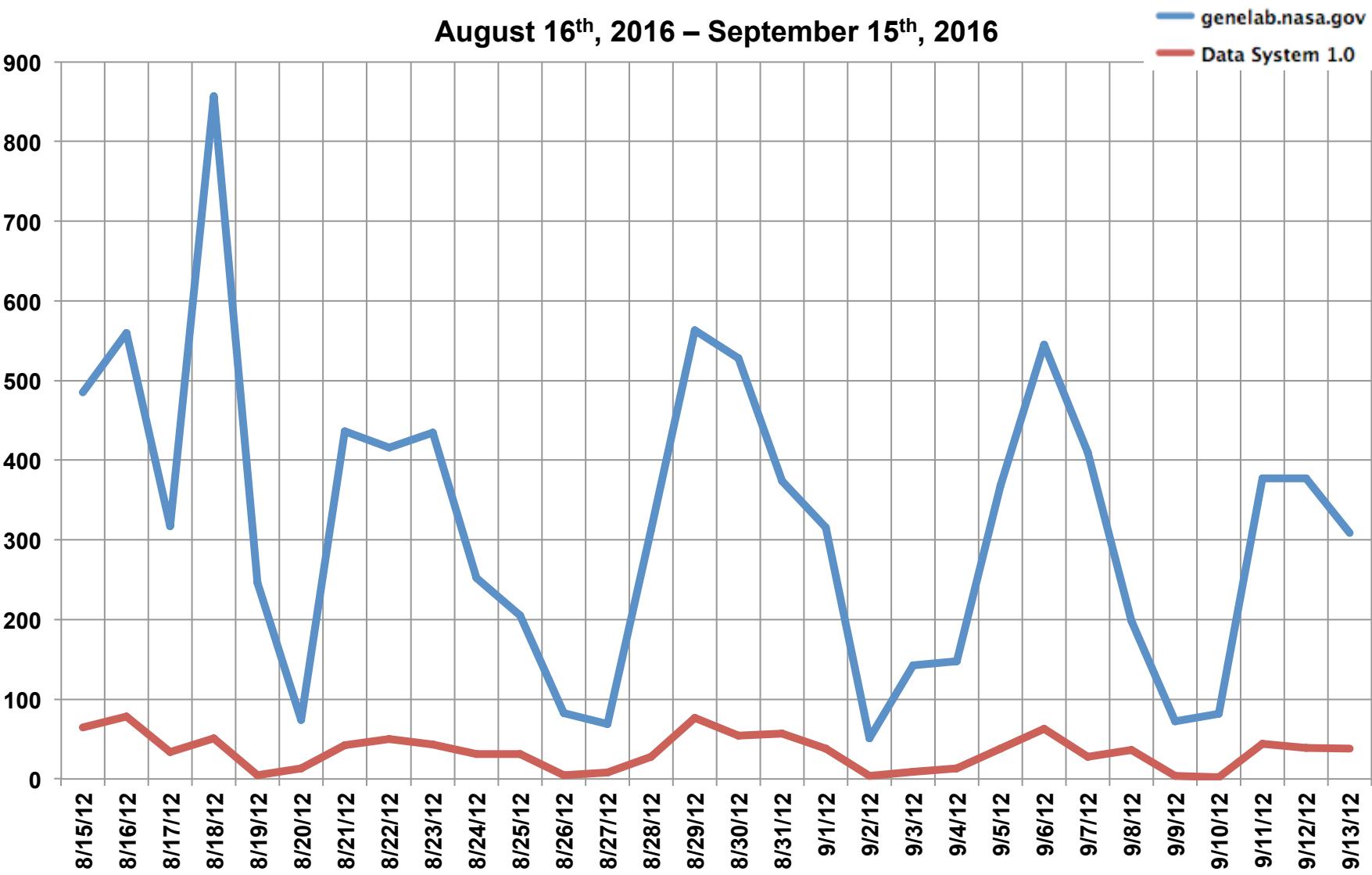
As of September 15th, 2016



| | | May 16 - Jun 15 | Jun 16 – Jul 15 | Jul 16 – Aug 15 | Aug 16 – Sep 15 |
|---|--------------------------|-----------------|-----------------|-----------------|-----------------|
| Monthly Pageviews | genelab.nasa.gov | 10,505 | 11,171 | 7,130 | 9,925 |
| | GeneLab DS 1.0 | 955 | 1,687 | 1,030 | 1,074 |
| Datafile Downloads - Totals | | | | | |
| | total count | 1,048 | 1,690 | 554 | 1,103 |
| | volume (GB) | 4,836 | 4,919 | 1,469 | 2,504 |
| Data File Downloads (by Species) | | | | | |
| | Arabidopsis thaliana | 399 | 437 | 35 | 187 |
| | Candida albicans | 4 | 10 | 0 | 3 |
| | Drosophila melanogaster | 24 | 56 | 7 | 22 |
| | Homo sapiens | 48 | 102 | 25 | 110 |
| | Mus musculus | 492 | 704 | 393 | 538 |
| | Pseudomonas aeruginosa | 6 | 20 | 2 | 6 |
| | Rattus norvegicus | 16 | 58 | 4 | 18 |
| | Saccharomyces cerevisiae | 6 | 45 | 8 | 13 |
| | Salmonella enterica | 4 | 9 | 0 | 4 |
| | Environmental Samples | 9 | 25 | 0 | 11 |
| | Bacillus subtilis | 5 | 12 | 0 | 4 |
| | Rhodospirillum rubrum | 13 | 30 | 0 | 8 |
| | Caenorhabditis elegans | 15 | 39 | 5 | 13 |
| | Escherichia coli | 3 | 9 | 1 | 3 |
| | Ceratopteris richardii | 4 | 10 | 3 | 2 |
| | Cellular organisms | – | 55 | 15 | 81 |
| | Bacillus | – | 19 | 14 | 17 |
| | Bacterial isolates | – | 19 | 32 | 28 |
| | Aspergillus fumigatus | – | 12 | 5 | 9 |
| | Brassica rapa | – | 10 | 3 | 23 |
| | Danio rerio | – | 9 | 2 | 3 |

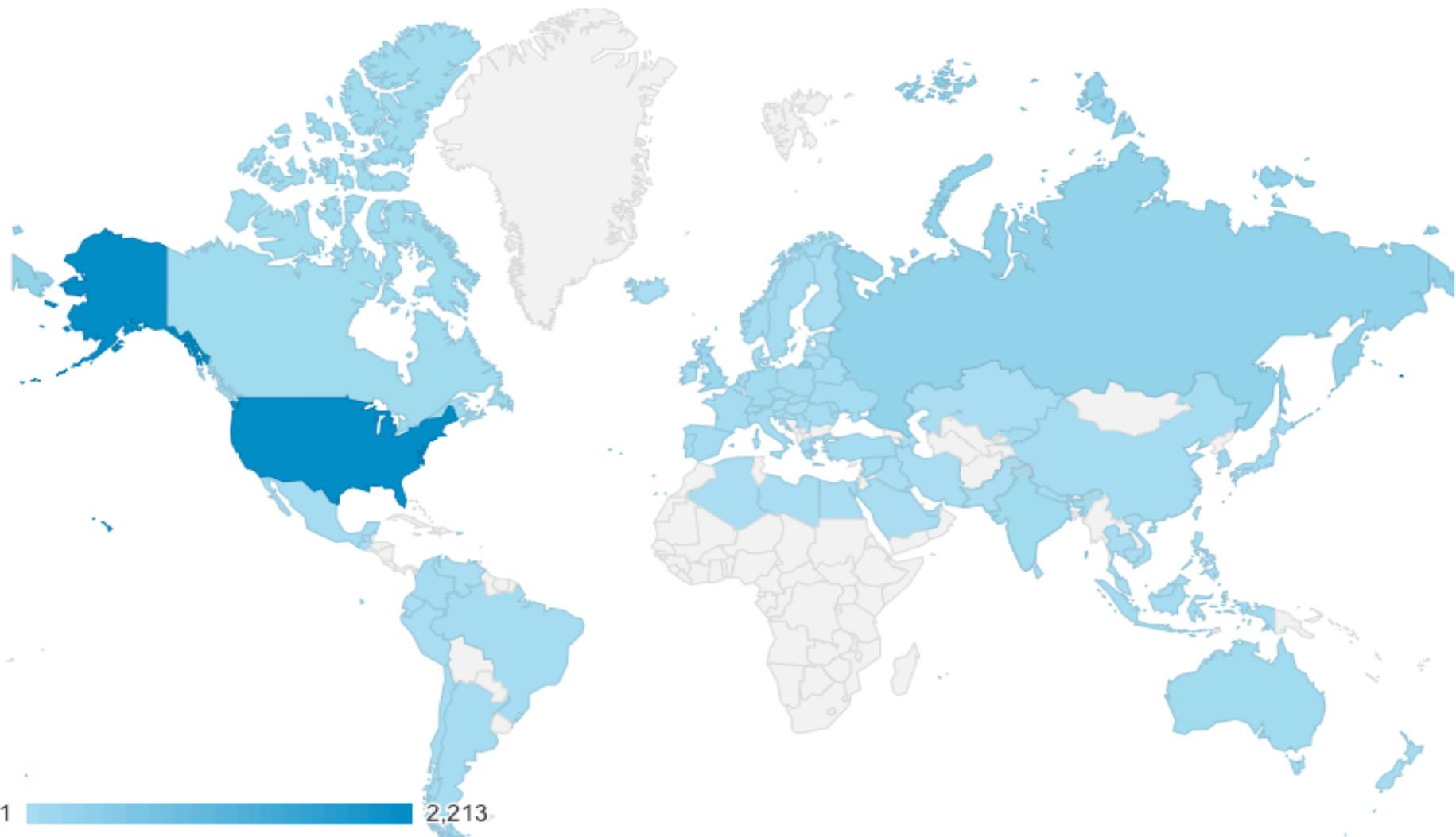


GeneLab Daily Page Views

August 16th, 2016 – September 15th, 2016

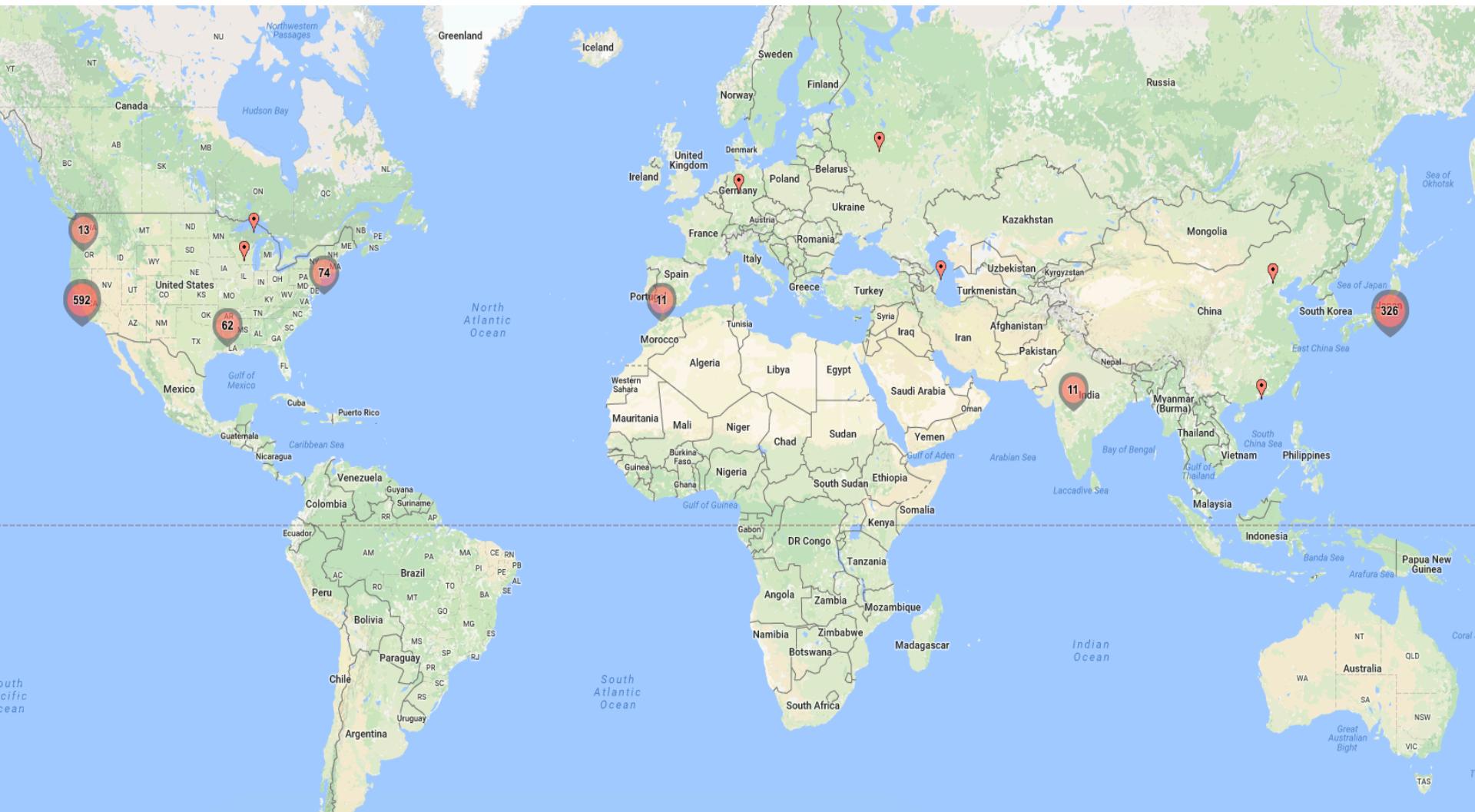


August 16th, 2016 – September 15th, 2016





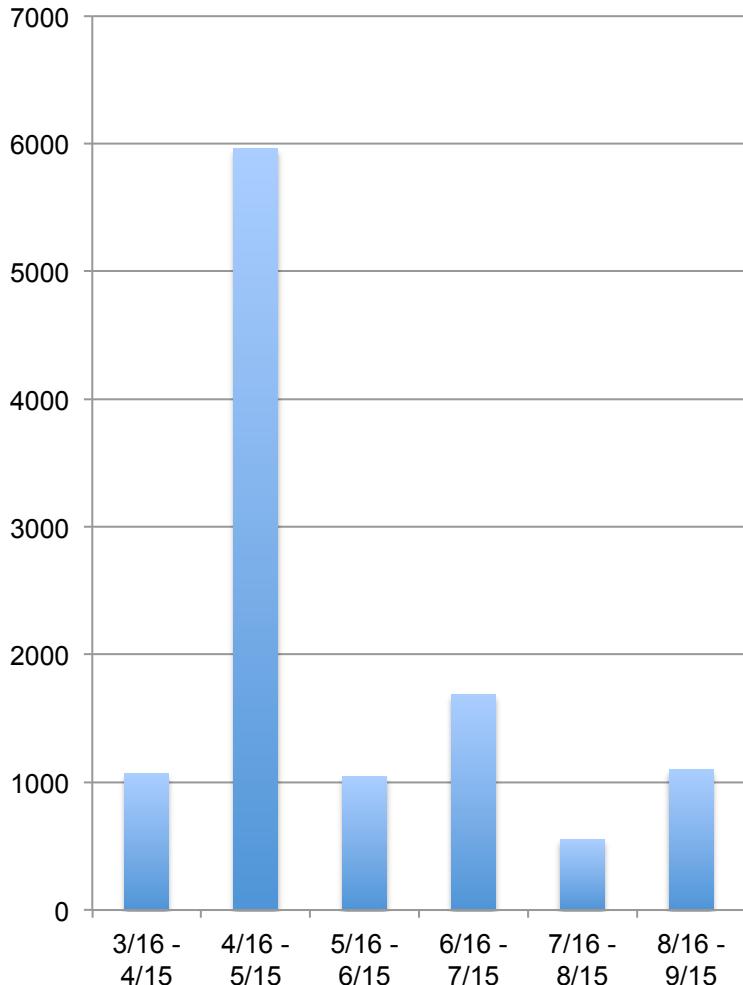
August 16th, 2016 – September 15th, 2016



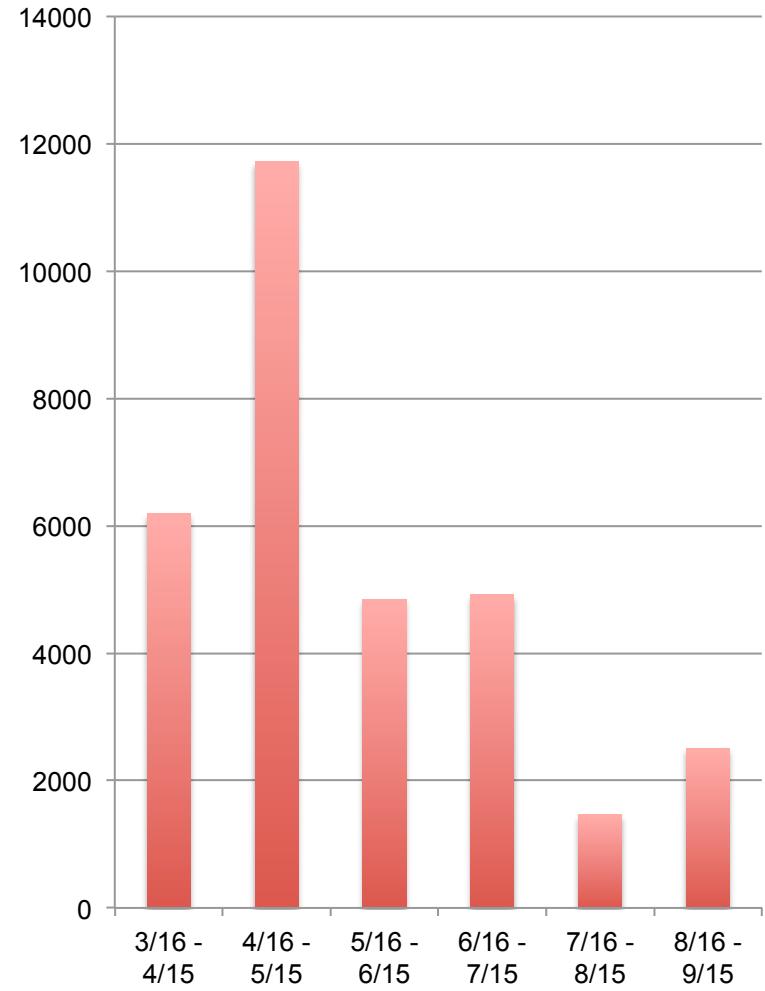


March 16th, 2016 - September 15th, 2016

Monthly File Downloads



Monthly Download Volume (GB)





Data File Downloads by Model Organisms



August 16th, 2016 to September 15th, 2016

- Mus musculus
- Arabidopsis thaliana
- Homo sapiens
- cellular organisms
- bacterial isolates
- Brassica rapa
- Drosophila melanogaster
- Rattus norvegicus
- Bacillus
- Saccharomyces cerevisiae
- Caenorhabditis elegans
- Environmental Samples
- Aspergillus fumigatus
- Rhodospirillum rubrum
- Pseudomonas aeruginosa
- Bacillus subtilis
- Salmonella enterica
- Danio rerio
- Escherichia coli
- Candida albicans

